

# 2801.001 Spring 2018 Homework 1

Martin Arienmughare – moa258

Madhur Bhattad – mb6854

Louis Guigo – lg2894

Mario Zhu – mz833

April 2018

**Submission instructions:** Groups of three to four, due 14 calendar days after the homework is posted, in electronic format. Submission must contain: (a) master answer sheet (typewritten or handwritten) with header containing your names and NetIDs, and (b) all relevant computer files such as code source, Excel files etc.

---

## Problem 1

You are short 100 at-the-money call option contracts on the S&P 500 index expiring in one year with a contract multiplier of 100. The current index level is 2600, the interest and dividend rates are zero.

- (a) Simulate the evolution of the index level at periods  $t$  as a geometric Brownian motion with volatility  $\sigma$  (free parameter) and calculate the corresponding call value, delta, gamma and theta using a fixed 20% implied volatility.

### Solution.

Please refer to the `HW1-GR1.ipynb` file for the code as well as the numerical values obtained, as well as plots showing the evolution of the stock price, the call price, the delta, the gamma and the theta (all these greeks being computed for one call option) accross time.

- (b) Then simulate your actual cumulative P&L when periodically delta-hedging your position (assuming you can trade the index as an asset) and compare it against the proxy formula on slide 11 for the following matrix of parameters:

### Solution.

Actual:

$\sigma \backslash \Delta t$	Monthly (12 per year)	Weekly (52 per year)	Daily (252 per year)
$\sigma = 25\%$	132756	127405	124553
$\sigma = 20\%$	131948	127201	124423
$\sigma = 15\%$	131196	127086	124408

Proxy:

$\sigma \backslash \Delta t$	Monthly (12 per year)	Weekly (52 per year)	Daily (252 per year)
$\sigma = 25\%$	153502	611820	2296570
$\sigma = 20\%$	119743	458939	1663050
$\sigma = 15\%$	85797	311566	1080117

- (c) Provide a statistical analysis of your results over 10,000 simulations, where one simulation is an entire index path.

**Solution.**

The above tables show the empirical means obtained after 10000 simulations of stock paths (with volatility uniformly drawn at each simulation between 0 and 0.35).

Here are the corresponding empirical standard deviations:

Actual:

$\sigma \backslash \Delta t$	Monthly (12 per year)	Weekly (52 per year)	Daily (252 per year)
$\sigma = 25\%$	47908	88435	107883
$\sigma = 20\%$	55320	93210	109882
$\sigma = 15\%$	64879	98548	112054

Proxy:

$\sigma \backslash \Delta t$	Monthly (12 per year)	Weekly (52 per year)	Daily (252 per year)
$\sigma = 25\%$	14911	166244	1156583
$\sigma = 20\%$	15569	148819	946564
$\sigma = 15\%$	15651	124899	713512

## Problem 2

Use your knowledge of how the VIX is calculated to show that the VIX is not the price of an investable asset. What about the square of the VIX?

### Solution.

The VIX is calculated from the fair short term variance of listed options on the S&P 500. The fair short term variance is in accordance with the fair strike of a variance swap calculated using price of a log contract with payoff as  $-\log(\frac{S(T)}{F})$ , where  $F$  is the price of forward of the underlying  $S$  with corresponding maturity.

Since VIX is just an index, the only possible way to invest in it is by replication. Unlike other indices like the S&P 500, replicating the VIX using S&P 500 options is not possible as the time value factor associated with the VIX is the square root of the time value factor that is associated with risk neutral discounting. (Since the formula for the VIX calculation as the risk neutral time value factor  $e^{rT}$  associated with the squared of the VIX).

An intuition of this exists from a mathematical standpoint. Assume we are in a Black-Scholes, and in particular that the dynamics of the stocks follow lognormal distributions:  $dS(t) = S(t)\mu dt + \sigma S(t)dW(t)$ . Assuming we can find a function that would play an analogue role as the logarithm in the variance swap replication. By Itô-Doeblin lemma, we would have  $df(S_t) = f'(S_t)dS_t + \frac{1}{2}f''(S_t)\sigma^2 S_t^2 dt = (\mu f'(S_t) + \frac{1}{2}f''(S_t)\sigma^2 S_t)S_t dt + f'(S_t)\sigma S_t dW_t$ . No matter what, this  $\sigma^2$  term will always be here. And it is exclusively a function of  $S$ , not of the instantaneous volatility because the payoffs of vanilla options is not dependent on the volatility. It's only a function of the stock price. Therefore we can't get rid of this random component, and there's no way to hedge it, i.e. there is no way to replicate it.

Therefore the only way to invest in the VIX is via VIX futures and VIX options.

VIX squared on the other hand, as described by the calculation of fair short term variance, is just  $\frac{2}{T}$  times the risk neutral expectation of the log contract. The VIX squared can be invested using variance swaps.

Conclusion: The VIX is not the price of an investable asset, whereas its square is.

### Problem 3

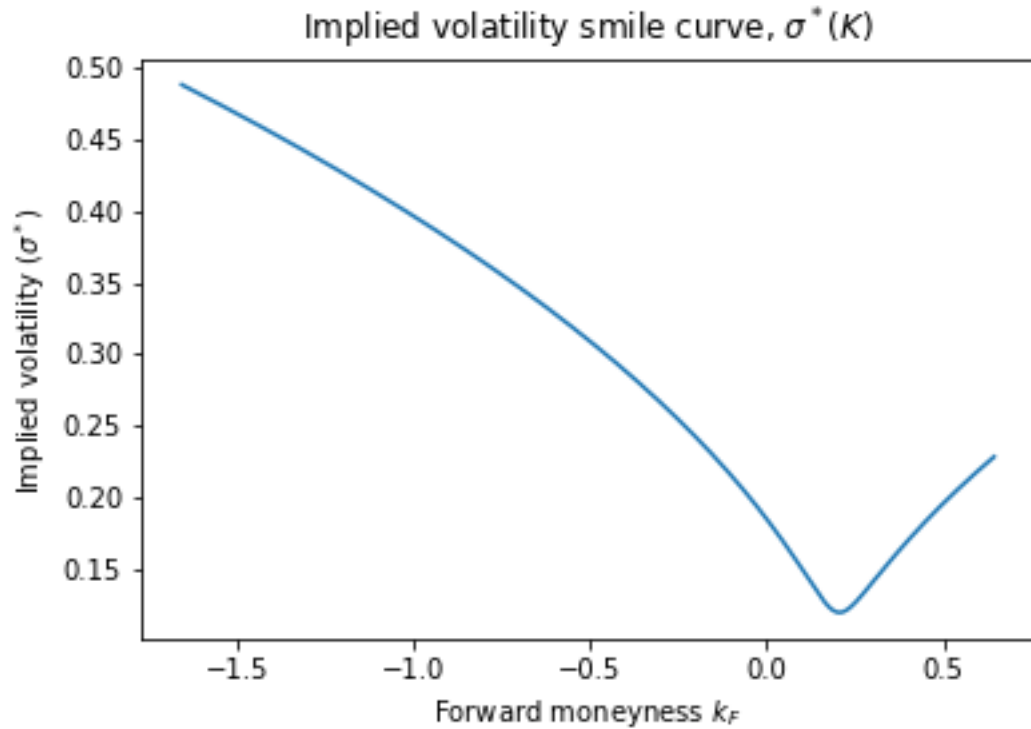
On March 29, 2018 the S&P 500 index (SPX) is 2611.53 and the 12-month VIX is 21.28. The implied volatility smile of SPX options expiring on March 15, 2019 is given as:

$$\sigma^* = \sqrt{a + b \left( \rho(x - m) + \sqrt{(x - m)^2 + s^2} \right)}$$

where  $a = 0.009$ ,  $b = 0.11$ ,  $\rho = 0.12$ ,  $m = 0.2$ ,  $s = 0.05$ ,  $x = \log \frac{K}{F}$  is log-moneyness and  $F = 2625.10$  is the forward price. The continuous interest rate is 2.09% p.a.

- (a) Draw the implied volatility smile curve for strikes  $500 \leq K \leq 5000$ .

**Solution.**



- (b) Calculate the fair strike  $K_{var}$  of a variance swap expiring on March 15, 2019 with the method of your choice. How close is your calculation to the 12-month VIX? Why is it not exactly the same?

**Solution.**

Applying *Eq.(5.3)* from the Advanced Equity Derivatives text to options data obtained from the CBOE on April 18 (3:28PM), resulted in the fair strike,  $K_{var} = 19.0$ . The strikes of the puts (calls) used were at most (least) the forward price. The results obtained are quite close to the 12-month VIX, with a percentage error of 11.8%. They are not exactly the same because variance is convex in volatility, and the difference between them is related to the volatility of the volatility.

## Problem 4

Consider a real symmetric matrix  $A$ . Show that  $\exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!}$  is symmetric positive-definite. Hint: Use spectral decomposition.

### Solution.

Let  $A \in S_n(\mathbb{R})$ . Let us prove that  $\exp(A) \in S_n^{++}(\mathbb{R})$ .

1. The linear application  $M \mapsto M^T$  is continuous on  $\mathcal{M}_n(\mathbb{R})$ , so if  $A$  is symmetric, we have  $\exp(A)^T = \exp(A)$ ;  $\exp(A)$  is symmetric.
2. Since  $A$  is symmetric, there exists an orthonormal basis of diagonalization of  $A$ , i.e.

$$A = P \text{Diag}(\lambda_1, \dots, \lambda_n) P^{-1}$$

with  $P \in O_n(\mathbb{R})$  and  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ . Taking the exponential of  $A$ , we therefore have:

$$\exp(A) = P \text{Diag}(e^{\lambda_1}, \dots, e^{\lambda_n}) P^{-1}$$

With this writing, we clearly see that all the eigenvalues of  $\exp(A)$  are strictly positive. Therefore  $\exp(A)$  is positive-definite.

With these two points,  $\exp(A)$  is a symmetric positive-definite matrix:  $\exp(A) \in S_n^{++}(\mathbb{R})$ .

## Problem 5

Let  $(\Omega, \mathcal{A}, \mathbb{P})$  be a probability space, and  $E$  the vector space of random variables with finite second moment and mean zero. Define  $\langle X, Y \rangle = \text{Cov}(X, Y)$  and, for any event  $A \in \mathcal{A}$ ,  $Z_A = I_A - \mathbb{P}(A)$  where  $I_A$  is the indicator variable of  $A$ .

- (a) Show that  $\langle \cdot, \cdot \rangle$  is an inner product on  $E$ . What is the induced norm?

### Solution.

Let us recall that, for  $X$  and  $Y$  random variables, we define the covariance by:

$$\text{Cov}(X, Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y)))$$

Relation from which we deduce by some algebra:

$$\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

Over  $E$ , it simplifies to  $\text{Cov}(X, Y) = \mathbb{E}(XY)$

Let us check that  $X, Y \mapsto \langle X, Y \rangle$  is an inner product on  $E$ , the vector space of real random variables with zero mean and finite second moment.

$\text{Cov}(\cdot, \cdot)$  is an application from  $E \times E$  to  $\mathbb{R}$ .

Let  $X, Y, Z \in E$  and  $\lambda \in \mathbb{R}$ .

- i) (*Symmetry*)  $\text{Cov}(X, Y) = \text{Cov}(Y, X)$ , clearly, using the symmetry in  $X, Y$  in the definition of the covariance.
- ii) (*Linearity*)  $\text{Cov}(\lambda X, Y) = \mathbb{E}((\lambda X - \mathbb{E}(\lambda X))(Y - \mathbb{E}(Y))) = \lambda \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) = \lambda \text{Cov}(X, Y)$ , by linearity of the expectation.  
 Besides:  $\text{Cov}(X + Z, Y) = \mathbb{E}((X + Z - \mathbb{E}(X + Z))(Y - \mathbb{E}(Y))) = \mathbb{E}((X - \mathbb{E}(X) + Z - \mathbb{E}(Z))(Y - \mathbb{E}(Y))) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) + \mathbb{E}((Z - \mathbb{E}(Z))(Y - \mathbb{E}(Y))) = \text{Cov}(X, Y) + \text{Cov}(Z, Y)$ .
- iii) (*Positive*)  $\text{Cov}(X, X) = \mathbb{E}((X - \mathbb{E}(X))^2) \geq 0$  by positivity of the random variable  $(X - \mathbb{E}(X))^2$  and monotony of the expectation.
- iv) (*Definite*)  $\text{Cov}(X, X) = 0$  iff  $(X - \mathbb{E}(X))^2 = 0$  since it is a positive random variable. This implies that  $X - \mathbb{E}(X) = 0$ , i.e.  $X = 0$  since  $\mathbb{E}(X) = 0$  by definition of  $E$ .

With these elements,  $\text{Cov}(\cdot, \cdot)$  is an inner product on  $E$ . The induced norm is:

$$\|X\| = \sqrt{\mathbb{E}(X^2)}$$

- (b) Show that  $|\rho(X, Y)| \leq 1$  (correlation coefficient).

**Solution.**

By definition,  $\rho(X, Y) = \frac{\langle X, Y \rangle}{\|X\| \|Y\|}$ . But  $\langle X, Y \rangle \leq \|X\| \|Y\|$  by Cauchy-Schwarz inequality, so  $\rho \leq 1$ .

- (c) Show that if  $X, Y \in E \setminus \{0\}$  are probabilistically independent then  $X, Y$  are linearly independent within  $E$ . Converse?

**Solution.**

Let us reason by contraposition. Let us assume that  $X, Y$  are not linearly independent, i.e. there exists  $\lambda \in \mathbb{R}$  such that  $Y = \lambda X$ . Necessarily,  $\lambda \neq 0$  since  $X, Y \neq 0$ .

So for  $x, y \in \mathbb{R}$ ,  $\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x, X = \frac{y}{\lambda})$ . Therefore,  $\mathbb{P}(X = x, Y = y) = 0$  if  $y \neq \frac{x}{\lambda}$  and  $\mathbb{P}(X = x)$  else. In that second case, we see that  $\mathbb{P}(X = x, Y = y) \neq \mathbb{P}(X = x)\mathbb{P}(Y = y) = \mathbb{P}(X = x)^2$ , unless  $X$  is constant equal to  $x$  which is impossible by assumption since  $X \neq 0$  and  $\mathbb{E}(X) = 0$  by definition of the vector space  $E$ .

Therefore,  $X$  and  $Y$  are not probabilistically independent.

This proves the direct implication over  $E$ .

The converse is false. As a counterexample, consider the random variable  $X$  taking values in the set  $\{-1, 0, 1\}$  with equal probabilities  $\frac{1}{3}$ , and take  $Y = X^2$ .

Clearly,  $X$  and  $Y$  are linearly independent (since the probability affected to the outcome  $-1$  is non zero in the law of  $X$ ). However,  $X$  and  $Y$  are probabilistically dependent since for instance  $\mathbb{P}(X = 0, Y = 1) = 0 \neq \mathbb{P}(X = 0) \cdot \mathbb{P}(Y = 1) = \frac{1}{3} \cdot \frac{2}{3}$ .

- (d) Let  $X, Y \in E \setminus \{0\}$ . What is the statistical interpretation of the orthogonal projection of  $Y$  on  $\text{Span}(X)$ ?

**Solution.**

The orthogonal projection of  $Y$  on the line  $\text{Span}(X)$  is exactly given by:

$$\Pi_{\text{Span}(X)}(Y) = \frac{\langle X, Y \rangle}{\|X\|^2} X = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} X$$

This orthogonal projection is a random variable linearly dependent with  $X$ , which has the same correlation with  $X$  as  $Y$  has. Additionally,  $\frac{\text{Cov}(X, Y)}{\text{Var}(X)}$  is the  $\lambda$  coefficient that minimizes the variance of any random variable  $Y - \lambda X$ .

In statistical terms, it means that if we wanted to predict  $Y$  using  $X$  with a linear model, i.e.  $\hat{Y} = \alpha X + \beta$ ,  $\frac{\text{Cov}(X, Y)}{\text{Var}(X)}$  would be the  $\alpha$  coefficient that minimizes the norm of the prediction error  $\epsilon = \mathbb{E}((\hat{Y} - Y)^2)$ .

Indeed the first order conditions of this simple minimization problem are:



$$\frac{\partial \epsilon}{\partial \alpha} = -2\mathbb{E}((\hat{Y} - Y)\frac{\partial \hat{Y}}{\partial \alpha}) \text{ and } \frac{\partial \epsilon}{\partial \beta} = -2\mathbb{E}((\hat{Y} - Y)\frac{\partial \hat{Y}}{\partial \beta})$$

which can be rewritten:

$$\mathbb{E}((\hat{Y} - Y)X) = 0 \text{ and } \mathbb{E}(\hat{Y} - Y) = 0$$

This means that the prediction error  $\hat{Y} - Y$  has to be orthogonal to  $X$  and that its expected value must be zero. Note that the first order conditions correspond to a minimum given the positive-definiteness of the Hessian.

Substituting  $\hat{Y} = \alpha X + \beta$ , we obtain  $\alpha = \frac{\text{Cov}(X, Y)}{\text{Var}(X)}$  and  $\beta = \mathbb{E}(Y) - \frac{\text{Cov}(X, Y)}{\text{Var}(X)}\mathbb{E}(X)$ . As stated above, it means that  $\frac{\text{Cov}(X, Y)}{\text{Var}(X)}X$  is the closest random variable (i.e. the one that minimizes the variance of the prediction error) to  $Y$  in  $\text{Span}(X)$ .

Besides, the prediction error is given by  $\epsilon = (1 - \rho^2)\text{Var}(Y)$ . If  $\rho = \pm 1$ , a perfect prediction can be made. When  $\rho = 0$ , the variance in the prediction is as large as the variation in  $Y$ , which implies that the predictor is very bad. For intermediate values of  $\rho$ , the predictor reduces the error.

- (e) Verify that  $Z_A \in E$  and calculate  $\langle Z_A, Z_B \rangle$  for any  $A, B \in \mathcal{A}$ . When is  $Z_A \perp Z_B$ ? Are  $Z_A, Z_{\bar{A}}$  linearly independent in  $E$ ?

**Solution.**

It is a fact that  $Z_A \in E$  for any  $A \in \mathcal{A}$  since this is a random variable with first and second moment respectively such that:  $\mathbb{E}(Z_A) = \mathbb{E}(I_A - \mathbb{P}(A)) = \mathbb{E}(I_A) - \mathbb{P}(A) = 0$  and  $\mathbb{E}(Z_A^2) = \mathbb{E}(I_A^2 - 2\mathbb{P}(A)I_A + \mathbb{P}(A)^2) = \mathbb{E}(I_A^2) - 2\mathbb{P}(A)\mathbb{E}(I_A) + \mathbb{P}(A)^2 = \mathbb{P}(A)(1 - \mathbb{P}(A))$ , so  $\mathbb{E}(Z_A^2) \leq \infty$  since  $I_A^2 = I_A$ .

Besides, for  $A, B \in \mathcal{A}$ ,  $\langle Z_A, Z_B \rangle = \langle I_A - \mathbb{P}(A), I_B - \mathbb{P}(B) \rangle = \langle I_A, I_B \rangle - \langle I_A, \mathbb{P}(B) \rangle - \langle \mathbb{P}(A), I_B \rangle + \langle \mathbb{P}(A), \mathbb{P}(B) \rangle = \langle I_A, I_B \rangle$

So,  $\langle Z_A, Z_B \rangle = \mathbb{E}((I_A - \mathbb{P}(A))(I_B - \mathbb{P}(B))) = \mathbb{E}(I_{A \cap B} - \mathbb{P}(B)I_A - \mathbb{P}(A)I_B + \mathbb{P}(A)\mathbb{P}(B))$ .

Id est,  $\langle Z_A, Z_B \rangle = \mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)$ .

We can see that  $Z_A \perp Z_B$  when  $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ , that is, when events  $A$  and  $B$  are stochastically independent.

$Z_A, Z_{\bar{A}}$  are not linearly independent in  $E$  since for any  $Z_{\bar{A}} = 1 - Z_A$ .

- (f) Suppose  $A, B \in \mathcal{A} \setminus \{\emptyset, \Omega\}$  are disjoint, and  $B \neq \bar{A}$ . Show that  $Z_A, Z_B$  are linearly independent.

**Solution.**

Let  $A, B \in \mathcal{A} \setminus \{\emptyset, \Omega\}$  such that  $A \cap B = \emptyset$ . Let  $\lambda, \mu \in \mathbb{R}$  such that  $\lambda Z_A + \mu Z_B = 0$ .

Since  $A \cap B = \emptyset$ , plugging in any two elementary events  $x \in A$  and  $y \in B$  yields:

$$\begin{cases} -\lambda\mathbb{P}(A) + \mu(1 - \mathbb{P}(B)) = 0 \\ \lambda(1 - \mathbb{P}(A)) - \mu\mathbb{P}(B) = 0 \end{cases}$$

Taking the difference between these equations, we obtain the condition:  $\lambda = \mu$ .

Therefore,  $\lambda(Z_A + Z_B) = 0$ . Since  $B \neq \bar{A}$ , this implies  $\lambda = \mu = 0$ .

Thus,  $Z_A$  and  $Z_B$  are linearly independent.

- (g) Suppose  $A \in \mathcal{A} \setminus \{\emptyset, \Omega\}$  and define  $B = \{\emptyset, A, \bar{A}, \Omega\} \subseteq \mathcal{A}$ . Let  $Y = E(X|B)$ . Verify that  $Y \in E$  and show that  $Y$  is the orthogonal projection of  $X$  on  $\text{Span}(Z_A)$ .

**Solution.**

Note that  $B$  is a  $\sigma$ -Algebra, therefore it makes sense to consider  $Y = \mathbb{E}(X|B)$ .

Let us recall that  $Y$  is defined as the almost surely unique random variable such that for any bounded and  $B$ -measurable random variable  $U$ ,  $\mathbb{E}(XU) = \mathbb{E}(YU)$ .

We have:  $\mathbb{E}(Y) = \mathbb{E}(\mathbb{E}(X|B)) = \mathbb{E}(X) = 0$  using the tower property of conditional expectation and the fact that  $X \in E$ .

Besides,  $\mathbb{E}(Y^2) = \mathbb{E}(\mathbb{E}(X|B)^2) \leq \mathbb{E}(\mathbb{E}(X|\sigma(X))^2) \leq \mathbb{E}(X^2) \leq \infty$  since  $X \in E$ .

Let us check that  $Y$  is the orthogonal projection of  $X$  on  $\text{Span}(Z_A)$ . For that matter, let us show that it is such that for all  $U \in \text{Span}(Z_A)$ ,  $\langle X - Y, U \rangle = 0$ .

Let  $U \in \text{Span}(Z_A)$ . Since  $Z_A$  is  $B$ -measurable,  $U$  is also  $B$ -measurable, and it follows that  $\mathbb{E}(UX|B) = U\mathbb{E}(X|B)$ .

Therefore:  $\mathbb{E}(U\mathbb{E}(X|B)) = \mathbb{E}(\mathbb{E}(UX|B)) = \mathbb{E}(UX)$  i.e.  $\langle U, Y \rangle = \langle U, X \rangle$ . That is exactly  $\langle U, Y - X \rangle = 0$ .

Since this is true for any  $U \in \text{Span}(Z_A)$ , we conclude that  $Y - X = \mathbb{E}(X|B) - X$  realizes the minimum distance from  $X$  to  $\text{Span}(Z_A)$ .

Since  $(E, \text{Cov}(\cdot, \cdot))$  is a Hilbert space, the minimum distance is realized by the orthogonal projection. Thus, we have  $Y = \Pi_{\text{Span}(Z_A)}(X)$ .

## Problem 6

Consider a vector space  $E$  equipped with a norm  $N(x)$ .

- (a) Show that  $N$  is Euclidean (i.e. induced by some inner product) if and only if  $N$  satisfies the parallelogram law:

$$N(x+y)^2 + N(x-y)^2 = 2 \cdot [N(x)^2 + N(y)^2]$$

Hint: Define the inner product through  $N$ .

### Solution.

If  $N$  is Euclidean, it derives from a certain inner product  $\langle \cdot, \cdot \rangle$ . And we have:

$$N(x+y)^2 + N(x-y)^2 = \langle x+y, x+y \rangle + \langle x-y, x-y \rangle = 2\langle x, x \rangle + 2\langle y, y \rangle + 2\langle x, y \rangle - 2\langle x, y \rangle$$

$$\text{Therefore, } N(x+y)^2 + N(x-y)^2 = 2 \cdot [N(x)^2 + N(y)^2]$$

Let's now assume that  $N$  checks the parallelogram law, and let us prove that the bilinear form  $\langle \cdot, \cdot \rangle$  defined by polarization as follows is an inner product over  $E$  from which  $N$  derives:

$$\langle x, y \rangle = \frac{1}{4} \cdot [N(x+y)^2 - N(x-y)^2]$$

We clearly have, by homogeneity,  $N(x) = \sqrt{\langle x, x \rangle}$ .

$\langle \cdot, \cdot \rangle$  is an application from  $E \times E$  to  $\mathbb{R}$ .

Let  $x, y, z \in E$  and  $\lambda \in \mathbb{R}$ .

- i) (*Symmetry*)  $\langle x, y \rangle = \langle y, x \rangle$ , clearly, using the symmetry in  $x, y$  in its definition.
- ii) (*Positive*)  $\langle x, x \rangle = \frac{1}{4} \cdot N(2x)^2 = N(x)^2 \geq 0$  since  $N(0) = 0$  by separation.
- iii) (*Definite*)  $\langle x, x \rangle = 0$  iff  $\frac{1}{4}N(x)^2 = 0$  iff  $x = 0$  by separation.
- iv) (*Linearity*) Observe that:

$$\langle x+y, z \rangle + \langle x-y, z \rangle = \frac{1}{4} [N(x+y+z)^2 - N(x+y-z)^2 + N(x-y+z)^2 - N(x-y-z)^2]$$

Therefore,

$$\langle x+y, z \rangle + \langle x-y, z \rangle = \frac{1}{4} [2 \cdot [N(x+y)^2 - N(z)^2] + 2 \cdot [N(x-y)^2 + N(z)^2]]$$

Thus,

$$\langle x+y, z \rangle + \langle x-y, z \rangle = 2 \cdot [N(x)^2 + N(y)^2]$$

Which can be rewritten:

$$\langle x+y, z \rangle + \langle x-y, z \rangle = 2 \cdot \langle x, y \rangle$$

From which we deduce, taking  $y = x$ :

$$\langle 2 \cdot x, y \rangle = 2 \cdot \langle x, y \rangle$$

By induction, we prove that this it is true for any power  $n$  that:

$$\langle 2^n \cdot x, y \rangle = 2^n \cdot \langle x, y \rangle$$

Taking  $u = \frac{1}{2}(x + y)$  and  $v = \frac{1}{2}(x - y)$ , we have:

$$\langle x + y, z \rangle = \langle 2 \cdot u, z \rangle = 2 \cdot \langle u, z \rangle = \langle u + v, z \rangle + \langle u - v, z \rangle = \langle x, z \rangle + \langle y, z \rangle$$

.

By induction we conclude that for any  $n \in \mathbb{N}$  (we've seen that it's true for  $n = 0$ ), and any  $x, y \in E$  we must have:

$$\langle nx, y \rangle = n \langle x, y \rangle$$

We can extend the linearity to  $\mathbb{Z}$  using  $\langle -x, y \rangle = -\langle x, y \rangle$ , and then to  $\mathbb{Q}$  using  $\langle x, y \rangle = n \cdot \langle \frac{1}{n}x, y \rangle$ .

We conclude that the linearity is true over  $\mathbb{R}$  using the well-known density of  $\mathbb{Q}$  in  $\mathbb{R}$  and the continuity of the application  $(x, y) \mapsto \frac{1}{4} \cdot [N(x + y)^2 - N(x - y)^2]$ , since the norm  $N$  is continuous.

Therefore, the linearity of  $\langle \cdot, \cdot \rangle$  is true.

Therefore  $\langle \cdot, \cdot \rangle$  is an inner product over  $E$  from which  $N$  derives since  $N(x) = \sqrt{\langle x, x \rangle}$ .

Conclusion:  $N$  is a Euclidean norm if and only if it checks the parallelogram identity.

(b) Which of the following are Euclidean norms on  $E = \mathbb{R}^n$ ?

$$N_p(x) = \sqrt[p]{\sum_{i=1}^n |x_i|^p}$$

$$Q_p(x) = \sqrt{N_2(x)^2 + \frac{1}{p} \sum_{i < j} x_i x_j}$$

where  $p \in [1, \infty)$ .

**Solution.**

Given the previous question, we just need to verify if these norms check the parallelogram identity to know if they are Euclidean.

For  $p = 2$ ,  $N_p$  is nothing but the so called Euclidean norm over  $\mathbb{R}^n$ , which as its name tells us, is Euclidean and derives from the canonical inner product of  $\mathbb{R}^n$ :  $\langle x, y \rangle = x^T y$ .

However, **for**  $p \geq 1$  **but**  $p \neq 2$ ,  $N_p$  is not Euclidean. As a counterexample, one can consider  $x = (1, 1, 0, \dots, 0)$  and  $y = (1, -1, 0, \dots, 0)$ . We have  $N_p(x) = N_p(y) = \sqrt[p]{2}$  but  $N_p(x+y) = N_p(x-y) = 2$ . So  $N_p^2(x+y) + N_p^2(x-y) = 8 \neq 2(N_p^2(x) + N_p^2(y)) = 4 \sqrt[p]{2}$ .

Since it doesn't check the parallelogram identity,  $N_p$  is not Euclidean **for**  $p \geq 1$  **but**  $p \neq 2$ .

$Q_p$  is not a norm for any  $p \in [1, \infty)$  since it doesn't check the triangle inequality: for  $x = (1, 0, \dots, 0)$  and  $y = (0, 1, 0, \dots, 0)$ , we have:

$$Q_p(x+y)^2 = 2^2 + \frac{1}{p} > Q_p(x)^2 + Q_p(y)^2 = 1^2 + 1^2 = 2$$

A fortiori, it can't be a Euclidean norm.

- (c) Define  $A^2(x, y) = [N(x) \cdot N(y)]^2 - [N(x+y)^2 - N(x-y)^2]^2/16$ . Show that if  $N$  is Euclidean then  $A^2(x, y) \geq 0$  and  $A(x+y, x-y) = 2 \cdot A(x, y)$ .

Geometric interpretation for  $E = \mathbb{R}^2$  ?

**Solution.**

The triangle and reverse triangle inequalities tell us that for any  $x, y \in E$ :

$$\begin{cases} N(x+y) \leq N(x) + N(y) \\ N(x-y) \geq |N(x) - N(y)| \end{cases}$$

Since both sides are positive, we can compose by the square function and obtain:

$$\begin{cases} N(x+y)^2 \leq (N(x) + N(y))^2 \\ N(x-y)^2 \geq (N(x) - N(y))^2 \end{cases}$$

Taking the negative version of the second inequality and adding them, we get:

$$N(x+y)^2 - N(x-y)^2 \leq (N(x) + N(y))^2 - (N(x) - N(y))^2$$

After simplification, the right hand-side is exactly  $4N(x)N(y)$  so we get:

$$\frac{1}{4}(N(x+y)^2 - N(x-y)^2) \leq N(x)N(y)$$

To prove  $A^2(x, y) \geq 0$ , we need to show that  $\frac{1}{16}(N(x+y)^2 - N(x-y)^2)^2 \leq (N(x)N(y))^2$ , i.e.  $\frac{1}{4}|N(x+y)^2 - N(x-y)^2| \leq N(x)N(y)$ .

Given the first inequality proved, we just need to show that

$$-\frac{1}{4}(N(x+y)^2 - N(x-y)^2) \leq N(x)N(y)$$

Given that  $N$  is Euclidean, this is equivalent to:

$$2(N(x)^2 + N(y)^2 - 2N(x+y)^2) \leq 4N(x)N(y)$$

Itself equivalent to

$$N(x+y)^2 \geq (N(x) - N(y))^2$$

Which is true by the second triangle inequality used above (it is its square).

Conclusion:

$$\frac{1}{4}|N(x+y)^2 - N(x-y)^2| \leq N(x)N(y)$$

i.e

$$\frac{1}{16}(N(x+y)^2 - N(x-y)^2)^2 \leq (N(x)N(y))^2$$

i.e.

$$A^2(x, y) \geq 0$$

Let us now prove that  $A(x+y, x-y) = 2 \cdot A(x, y)$  (well defined with the first point).

We have:  $4A^2(x, y) = 4((N(x)N(y))^2 - \frac{1}{16}(N(x+y)^2 - N(x-y)^2)^2)$ .

Using the parallelogram identity, we substitute:  $N^2(x), N^2(y), N^2(x+y), N^2(x-y) \leftarrow \frac{1}{2}(N^2(x+y) + N^2(x-y) - N^2(y)), \frac{1}{2}(N^2(x+y) + N^2(x-y) - N^2(x)), 2(N^2(x) + N^2(y)) - N^2(x-y), 2(N^2(x) + N^2(y)) - N^2(x+y)$ .

After draft calculations not reported here, we get exactly

$$A^2(x+y, x-y) = 4 \cdot A^2(x, y)$$

Taking the square root,

$$A(x+y, x-y) = 2 \cdot A(x, y)$$

In  $\mathbb{R}^2$ ,  $A^2$  represents the difference of squared surface between a rectangle with sides  $N(x)$  and  $N(y)$ , and a rectangle with sides  $\frac{N(x+y)}{2}$  and  $\frac{N(x-y)}{2}$ .

So the last identity means that the difference of squared surface between a rectangle with sides  $N(x + y)$  and  $N(x - y)$ , and a rectangle with sides  $2N(x)$  and  $2N(y)$  is equal to twice the difference of squared surface between a rectangle with sides  $N(x)$  and  $N(y)$ , and a rectangle with sides  $\frac{N(x+y)}{2}$  and  $\frac{N(x-y)}{2}$ .

Quite unsatisfying interpretation...